

Real-Time Adaptive Aircraft Scheduling

Stephan E. Kolitz and Mostafa Terrab

Charles Stark Draper Laboratory, Inc.
Cambridge, Massachusetts

Prepared for
Ames Research Center
CONTRACT NAS2-12419
June 1990



National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California 94035-1000

TABLE OF CONTENTS

Executive Summary	1
1. Introduction	3
1.1 Background	4
1.2 Problem Description	15
2. Deterministic Case	21
2.1 Standard Formulations	21
2.1.1 Mathematical Model	21
2.1.2 Minimum Cost Flow	24
2.1.3 Assignment Problem	26
2.2 The Fast Algorithm	28
2.2.1 Regular Costs	28
2.2.2 Notation, Pseudo-Code and a Flow-Chart	30
2.2.3 A Simple Example	34
2.3 Adaptive Considerations	38
3. Probabilistic Case	39
3.1 Problem Description	39
3.2 The MMR Algorithm	42
3.2.1 Cost functions, Pseudo-Code and a Flow-Chart	42
3.2.2 A Simple Example	47
3.3 Adaptive Considerations and Future Directions	54
4. Numerical Examples	57
4.1 Introduction	57
4.2 Deterministic Case	58
4.3 Probabilistic Case	61
5. References	65
6. Acknowledgements	67

EXECUTIVE SUMMARY

In recent years, air traffic systems in the United States and Europe have been experiencing more and more network-wide congestion. It is estimated that direct operating costs for U.S. commercial airlines due to delays amounted to \$2 billion in 1986 and a recent article in the Economist (November 1989) estimated \$8 billion for yearly delay costs for European airlines.

One of the most important functions of any air traffic management system is the assignment of "ground-holding" times to flights, i.e., the determination of whether and by how much the take-off of a particular aircraft headed for a congested part of the ATC system should be postponed in order to reduce the likelihood and extent of airborne delays.

We present an analysis of the fundamental case in which flights from many destinations must be scheduled for arrival at a single congested airport; the formulation is also useful in scheduling the landing of airborne flights within the extended terminal area. We describe a set of approaches for addressing a deterministic and a probabilistic version of this problem.

For the deterministic case, where airport capacities are known and fixed, we develop several models with associated low-order polynomial-time algorithms. For general delay cost functions, these algorithms find an optimal solution. Under a particular natural assumption regarding the delay cost function, we develop an extremely fast ($O(n \ln n)$) algorithm.

For the probabilistic case, using an estimated probability distribution of airport capacities, we develop a model with associated low-order polynomial-time heuristic algorithm with useful properties.

We test the models on data simulated from actual 1987 Logan Airport (Boston) data. Results indicate that improvement in costs in the range of 30% to 50% are achievable, even in equitable but sub-optimal cases; far larger savings are possible using policies closer to optimal.

The central conclusion of the work is the following: **if there are delays in the system due to capacity problems, then there is a large payoff from using scheduling policies other than first come, first served.**

1. INTRODUCTION

A general aim of research efforts in the past several years in the area of air traffic control (ATC) automation has been to develop and test systems designed to reduce controller peak work, to improve the man-machine interface, and to achieve better integration of information across system interfaces. For near-term implementation, automation of selected ATC functions in the extended terminal area from 150 nautical miles to touchdown using a time based framework has been studied. (See Tobias[1985] and Alcabin[1985].) Algorithms have been developed and procedures tested in real-time simulation studies for handling time assignment of new arrivals, correction for time errors as aircraft progress through the terminal area, missed approaches, and runway closures (Tobias[1986]). Results are very encouraging that the time-based framework and associated algorithms form a sound basis for future incremental improvement in the ATC system. All of the time-based schedule planning, however, carries over the first come, first served (FCFS) server discipline that is embodied in current manual control practice. This is probably necessary for consideration of near-term implementation in the ATC system, but unnecessarily limiting for the far-term.

The original objective of the research was to work within a proposed scheme for time-based control in the extended terminal area but without the FCFS restriction. Once the FCFS assumption was lifted, however, it became clear that real-time adaptive aircraft scheduling problems, (both system-wide and in the extended terminal area) must be understood before any other work could proceed productively. This report is a rigorous look at a particular variant of real-time adaptive aircraft scheduling, the "Ground Holding Policy Problem."

The algorithms developed in this report can be run very quickly, hence “real-time;” they can also be initialized with the observed current state of the system (which flights are on the ground, which are in the air, delays, etc.) giving results that depend upon the evolution of the system during a day, hence “adaptive.” Also, “ground holds” can be understood to be not only actual holds on the ground, but can be used generically to refer to airborne holds in some situations.

1.1 Background

In recent years, air traffic systems in the United States and Europe have been experiencing more and more network-wide congestion. It is estimated that direct operating costs for U.S. commercial airlines due to delays amounted to \$2 billion in 1986 and a recent article in the Economist (November 1989) estimated \$8 billion for yearly delay costs for European airlines. Moreover, the impact of these delays goes far beyond these direct costs to airlines, since a large proportion of passengers are also affected. Approximately 70% of all emplanements and deplanements in the U.S. take place in only 22 major airports (the 22 “pacing” airports in FAA terminology) most of which constitute the principal bottlenecks of the air traffic network (90% to 95% of delays can be attributed to congestion in these pacing airports). This situation is not likely to improve in the near future as demand for air transportation to and from major metropolitan areas is forecast to continue to grow. Furthermore, the problem is exacerbated by the use of the “hub and spoke” scheduling system by almost all major airlines. This system consists of using, for reasons of operational and economic efficiency, a small number of designated airports as transit points for a large proportion of all flights of a given airline, thereby generating a large number of operations at these airports.

The existing literature dealing with the ATC flow management problem is rather limited. One reason for this may be that network-wide ATC delays were not felt until 1986. Odoni [1987] has the first systematic description of the flow management problem (FMP). The paper provides a mathematical formulation of the FMP and points to the need for analytical and algorithmic approaches to this problem. It also contains a discussion of the interplay between the technical and policy aspects of the FMP.

We will refer to the single airport version of the FMP as the Ground Holding Policy Problem (GHPP) throughout this report. To our knowledge the only algorithmic approach dealing specifically with the GHPP is a paper by Andreatta and Romanin-Jacur [1987]. In this paper, the authors propose a dynamic programming approach to solve a probabilistic version of the GHPP for which it is assumed that a destination airport experiences congestion during a single time period.

Even though algorithmic approaches are almost nonexistent there have been a number of computer-simulation based approaches to the FMP. The Federal Aviation Administration (FAA) currently uses such a system to determine ground holds.

The FAA is responsible for initiating and coordinating ground-holding strategies in the U.S. air traffic system. The FAA operates a Central Flow Control Facility (CFCF) in Washington, D.C. for this purpose; the CFCF is equipped with state-of-the-art information gathering capabilities through which it obtains access to regional and local weather data and forecasts as well as up-to-the-minute information on the status of virtually all airborne traffic in the U.S.

Part of the information the CFCF obtains from local control centers is a estimate of airport capacities for a given time period. This information is then used to determine ground holds for individual flights by:

- estimating (through what is essentially a deterministic simulation model) the air-delay that could be expected for each flight if it left at its scheduled departure time.

Figure 1-1 illustrates this process.

- setting these ground holds equal to the "expected" airborne delay if the latter is not below a given threshold (typically the "threshold" may be set at 15 or 20 minutes).

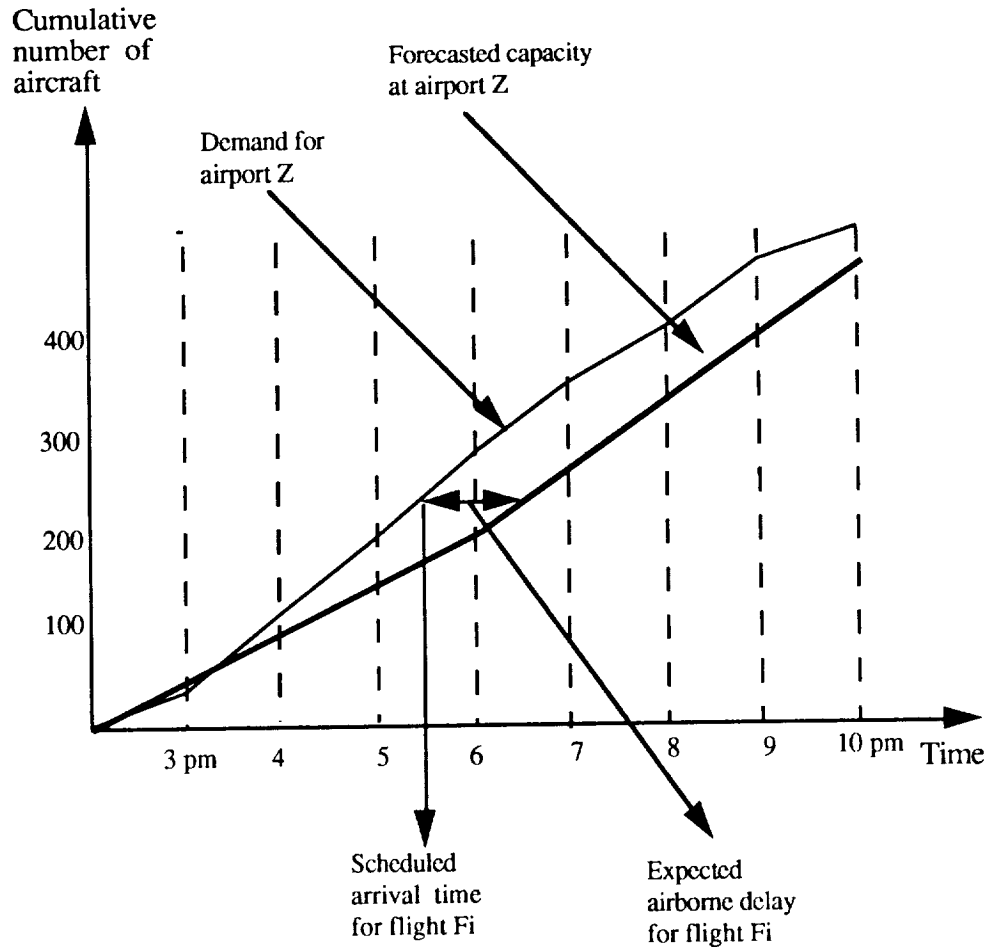


Figure 1-1

This process can therefore be described as assigning flights to available capacity on a first-come first-served (FCFS) basis (i.e. flights with earlier scheduled arrival times are given priority) without considering any uncertainties in the capacity estimate.

Assuming the capacity estimate used in this process is indeed accurate, this approach would minimize the total delay costs if all flights using airport Z had identical delay costs per unit of time (since it minimizes total aircraft-delay). It does not, however, consider

cost functions that reflect different operating costs for different sizes of aircraft. It also does not consider explicitly the uncertainty concerning the capacity estimate itself.

The air traffic network congestion problem can be addressed according to different time spans:

- Long-term approaches include the construction of additional airports (which typically requires 10 to 15 years from conception to operation), improved Air Traffic Control (ATC) technologies, additional runways at existing airports, and use of larger aircraft. Although several such airport improvement programs are under way in the U.S. and abroad, these approaches are generally very costly and may often be difficult to implement due to a lack of public support. This is particularly true for airport extension projects; these projects are usually needed in larger metropolitan areas where demand for air travel is high, but where resistance to such projects is also high due to their impacts on local communities.

- Medium-term approaches (6 months to a few years) are mostly administrative or economic in nature. They try to alleviate congestion by modifying the temporal pattern of aircraft flow through the ATC network, for example by imposing different user charges at different times of day or putting pressure on airlines to modify some of their scheduling practices, such as the use of the hub-and-spoke system, that tend to concentrate flights temporally and geographically.

- Short-term approaches have to deal with a given schedule and network capacity and are intended to mitigate the effects of unavoidable congestion through the control of the flow of aircraft. The time horizon for such flow control can vary from a few minutes to a whole day and the purpose is to best match the flow of aircraft to available capacity throughout the time horizon.

This report focuses on a particular version of these short-term problems that consists of considering ground holds on some flights before departure. While the emphasis and terminology will relate to ground holds before departure, much of the analysis can be used for airborne flights, including those in a 150 nautical miles neighborhood of a terminal.

In order to describe the FMP we use a network representation of the ATC system that distinguishes four types of elements, as shown in Figure 1-2:

- Airway elements are represented by arcs in the network. These correspond to the physical paths that aircraft use.
- Airport elements are the sources and sinks for air traffic in the system; they are represented by nodes in the network.
- Waypoint elements are also represented by nodes in the network. They correspond to points in the network where airways intersect.
- Sector elements are defined to be a set of waypoints and segments of airways that is treated as a unit for air traffic control purposes.

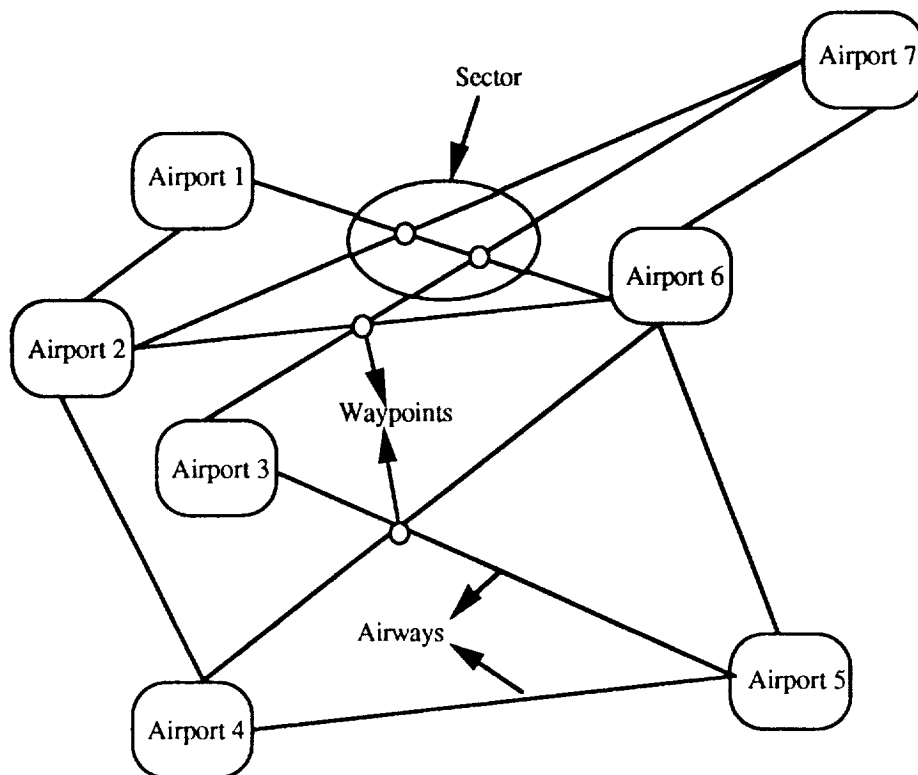


Figure 1-2

Congestion can occur at any one of the elements of the network of Figure 1-2 when the capacity of these elements is reduced. If the capacity of each one of the elements of the network were known and did not change with time, there might be no delays in the ATC system since the flows could be adjusted to match exactly these known capacities. Delays occur because these capacities, in particular the airport capacities, can be greatly affected by conditions that change over time and may be difficult to predict. Airports constitute the principal bottlenecks of the ATC system. Primarily, weather conditions (visibility, precipitation, wind, cloud ceiling, etc...) determine the capacity of a given airport (in terms of the number of landings and take-offs that can occur during a given time span) since they affect minimum landing separation rules and can dictate the

runway configuration to be used. The reductions in airport capacity can be as high as 50% in some extreme weather conditions. Figure 1-3 shows the runway capacity profile for Logan Airport (Boston) for the year 1987; it shows in particular that, for that year, the capacity of Logan Airport was reduced from a maximum of 130 operations per hour to less than 100 operations per hour more than 20% of the time, with a reduction to less than half that maximum capacity (60 operations per hour) happening 15% of the time.

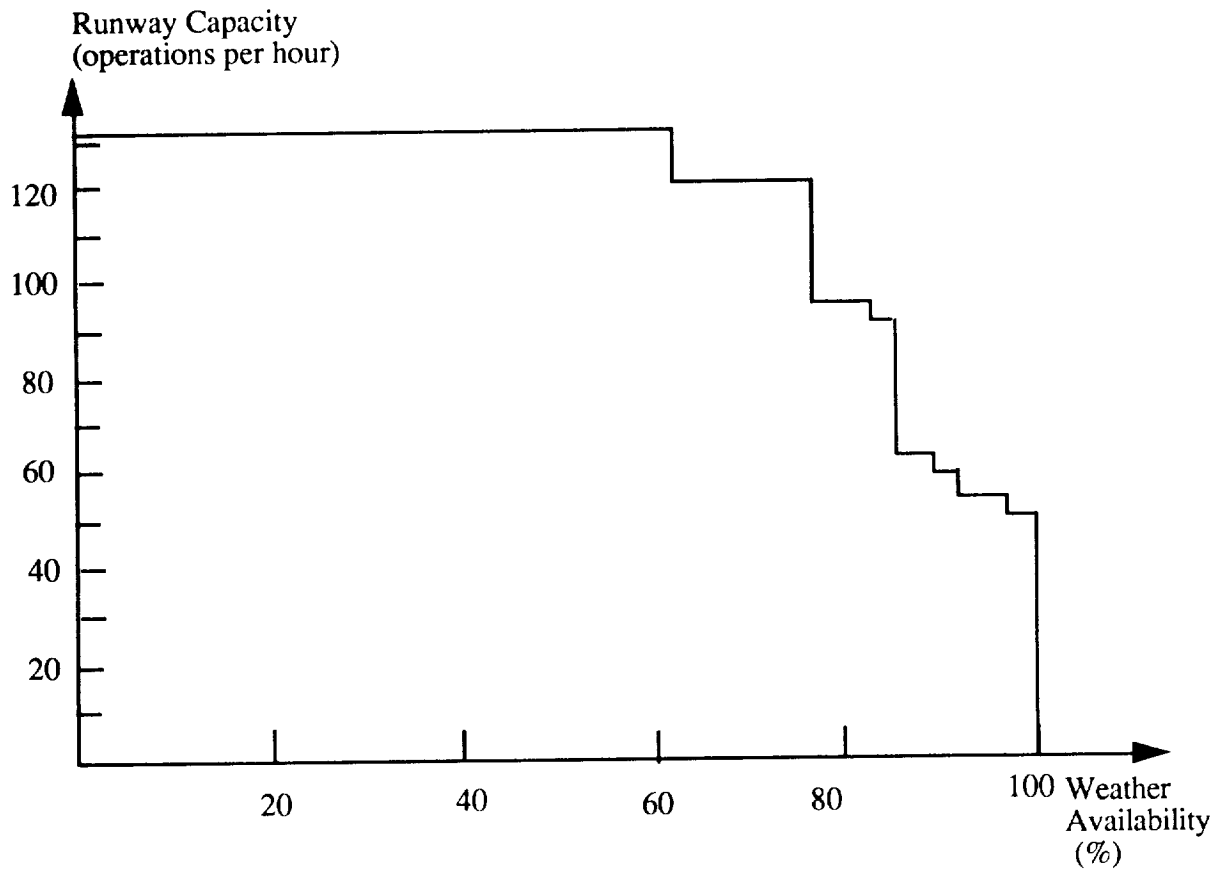


Figure 1-3

There is a need to adjust the aircraft flows through the various elements of the ATC network on a short term basis using available information and estimates concerning the capacities of airports and other elements of the ATC network. The FMP can then be defined as the problem of adjusting flows in the ATC network so as to minimize the cost of delays given the available information concerning the (present and future) status of the elements of the ATC network. Odoni [1987] distinguishes two types of such flow management actions. “Tactical” actions are to be exercised when the aircraft is already airborne and include:

- High altitude holdings, path stretching manoeuvres, or modifying en-route flight plans in order to avoid costlier low altitude delays.
- The control of en-route speeds to time the arrival of an aircraft at a congested point of the ATC network.
- The sequencing of aircraft for landing to maximize runway acceptance rate.

These actions can help control the flow of aircraft through specific areas of the air traffic network but are limited in the extent of control they permit by the very fact that the aircraft is airborne and is therefore subject to fuel and safety constraints.

The other type of action is a “strategic” type of action with greater potential for regulating aircraft flows. It includes:

- The modification of flight plans of some flights before take-off in order to bypass congested areas of the network.
- Delaying the departure time of some flights. These delays are referred to as “gate holds” or “ground holds” and correspond to delaying the actual departure time of an

aircraft beyond its scheduled take-off time. These delays are to be taken before the aircraft starts its engines on the apron area (either at the gate or in a remote parking area) even if the aircraft is otherwise ready to taxi to the runway.

The first type of strategic action also has a limited potential since there are clearly limits to the scope of modifications of flight plans (e.g., fuel constraints). The second type of action, on the other hand, provides greater flexibility in adjusting flows and can lead to greater savings in delay costs for the following two reasons:

- First, because our ability to control and regulate aircraft flows is greatly increased since we do not have to deal with some of the constraints (e.g. fuel constraint) of airborne control.

- Second, because important savings in delay costs as well as improved safety can be expected if we are able to absorb some of the delays on the ground. Savings in delay costs can be expected because ground-holding delays are less costly than airborne delays that involve fuel consumption as well as depreciation and maintenance costs. The safety issue is fresh in our minds, with the very recent crash of the Avianca plane on Long Island. Since, in general, it is not possible to predict airport capacities exactly at the times of departure of aircraft, the basic trade-off is between the cost of airborne delays that result from optimistic strategies, which impose little ground holds, and the costs of ground holds from pessimistic strategies, which impose excessive ground holds and may result in unused capacity.

This type of strategic flow management action should therefore be considered as the essential component of a flow management system; it should, nevertheless, be complemented by the tactical actions mentioned above. The models developed in this report can be used to evaluate and help decide both tactical and strategic actions. It is

also important to note that the FMP is not a short-term problem in the sense that it is to be addressed only until the longer term issues are resolved. It is, in fact, to be considered a permanent part of an efficient ATC system. The reason for this is that, as we have argued above, the FMP is motivated by the fact that the capacity conditions in the ATC system can change over time and may be difficult to predict. This situation is not likely to be greatly affected by longer term actions (except, maybe, in terms of improvement in weather prediction technology) and there will always be a need to adjust flows in the short term to match available capacity for an efficient use of the ATC network.

Although the above classification of methods for tackling the FMP is useful for decomposing the problem, we should keep in mind that the boundaries between tactical versus strategic flow management are not clearly defined in practical situations. There are some obvious interdependencies, such as, for example, the fact that airport capacity depends not only on weather conditions but also on tactical actions such as sequencing of landing aircraft. A complete flow management system has to deal with interfacing the various approaches.

As is the case with airport capacities, capacities of airways, waypoints and sectors (in terms of the number of aircraft that can traverse these elements per unit time or occupy them at any given time) can also be reduced. In most cases, however, the remaining capacities for these elements are greater than the actual traffic flows that use them. This suggests considering, as a first approximation, strategic flow actions when we assume that the only capacitated elements of the air traffic network are the airports. The resulting version of the FMP that deals exclusively with the trade-off between ground-holds and airborne delays for a network in which airports are the only congested elements is called the “generic FMP” in Odoni [1987].

The main focus of this report is a fundamental case of this problem, the GHPP; we must decide on ground holds for flights from many destinations scheduled for arrival at a single congested airport. There are several reasons for looking at the single arrival airport case:

- There are instances for which this is a reasonable formulation. This is true when only one airport of the network is expected to be congested. It is also applicable to cases for which several airports are expected to be congested but there is little traffic between these congested airports (or if the congestion is not affecting the inter-airport traffic on both ends simultaneously).
- It features and focuses on the most important aspect of the strategic FMP: the trade-off between ground-holding delays and airborne delays.
- Solution methods for the full network case can be based on solution methods for the single airport case.

1.2 Problem Description

The GHPP is concerned with timing the arrival of aircraft coming from several origin airports into a single arrival airport, Airport Z, and is described, referring to Figure 1-4, as follows:

- (1) We consider arrival operations at a given airport Z during a time interval $[0, T]$ for which we expect some amount of congestion.

(2) We are given a complete list F_1, F_2, \dots, F_N of flights departing from a number of origin airports and scheduled to land at airport Z during $[0, T]$. For each flight F_i , we know the scheduled departure and arrival times of F_i ; furthermore we assume that the travel times are deterministic and known in advance.

(3) The time period $[0, T]$ is discretized into P time periods T_1, T_2, \dots, T_P and we are given a estimate of the observed values of the random variables K_1, K_2, \dots, K_P , the corresponding arrival capacities of airport Z. Thus K_j is the maximum number of landings that can take place at airport Z during T_j .

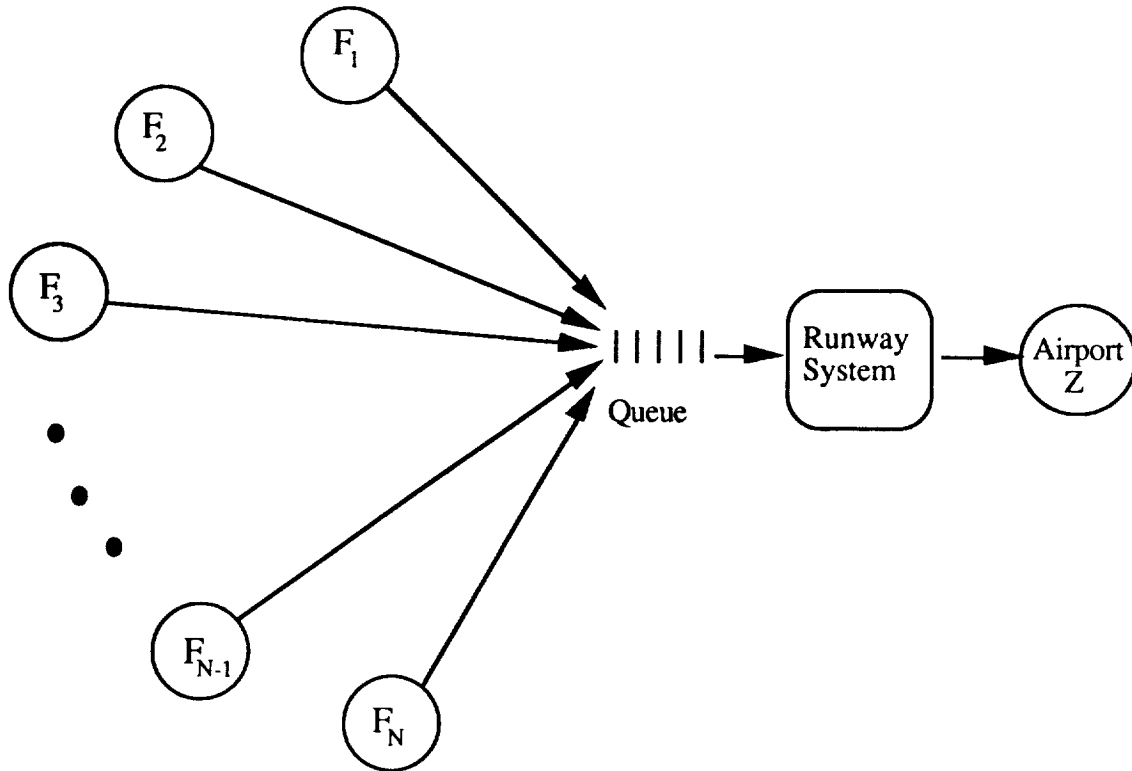


Figure 1-4

The GHPP is then defined as the problem of finding, for each flight F_i , the optimal amount of ground hold to be imposed on flight F_i so that the overall expected total delay cost (ground + airborne queueing delays) is minimized. A solution to the GHPP will be referred to as a Ground Holding Policy (GHP) in future discussions.

If the queueing delays were predictable, we would take them entirely as ground delays before departure so as to minimize costs. Because of the uncertainty concerning the capacity of airport Z during $[0, T]$, however, we cannot predict the amount of queueing delays that each flight would incur if it left on time. The GHPP is therefore concerned with finding the amount of ground delay to impose on each flight so as to strike an optimal balance between the costs of these (known) ground delays and future (unknown) airborne delay costs based on available information concerning the capacity of airport Z. The most general version of the GHPP has therefore the following characteristics:

(a) It is a stochastic problem. There are many circumstances for which it is not possible to exactly predict the capacity of an airport even a few hours in advance. The reason is that capacity depends strongly on local weather conditions which are subject to high uncertainty. In these circumstances we may have to settle for a probabilistic capacity estimate. The object of the GHPP is then to try to reach an optimal balance between (present-known) ground delays and (future-expected) airborne delays.

(b) It is a dynamic problem. It is possible that the probability distributions of the capacities are not stationary; that is, they can change over time.

(c) It is a combinatorial problem. Because the individual delay costs for aircraft can be very different, it is not enough to think of this problem solely in terms of controlling the flow of aircraft arriving at airport Z; we have to determine the optimal composition of these flows as well. The combinatorial aspect is evident if we consider the case for which we can predict the capacity of airport Z exactly (deterministic capacity estimate); we can consider this situation as one for which the stochastic and dynamic aspects have been resolved but it is obvious that large cost-savings can be expected from assigning individual aircraft to available capacity in an optimal manner.

The GHPP assumes that the arrival capacities K_1, K_2, \dots, K_p of airport Z are exogenous. However, there is however an interaction between the actual capacities and the GHP resulting from solving the GHPP. The reason for this is that the acceptance rate of a given runway configuration during a given time period depends on the aircraft mix during that time period through a set of minimum separation rules between aircraft types. The GHPP also assumes that we can determine arrival capacities separately from departure operations at airport Z. In some cases, when arrivals and departures share the same runways, there is an interaction between both types of operations. It is assumed here that the split of total capacity between arrivals and departures is known and remains constant for each time period T_i .

Finally, the model assumes implicitly that there is no departure congestion at the airports of origin so that a given GHP can be implemented regardless of conditions at these origin airports. Including such considerations into the model is in the scope of multi-airport formulations (the FMP).

The probabilistic aspects of this problem are difficult to deal with. The approach taken in this report is to decompose the problem. We first develop solution methods for

the deterministic version, where the arrival capacities are fixed and known values, resulting in a combinatorial problem. Then we focus on the probabilistic version of the problem, where the capacities are random variables, in which determining ground holds is based on a estimation of the capacities.

2. DETERMINISTIC CASE

This section is concerned with the deterministic version of the GHPP which assumes that future capacity at airport Z can be predicted *exactly*. This assumption reduces the GHPP to a purely combinatorial problem for which standard solution methods are available. One of the motivations for looking at the deterministic case is that there are cases for which it is reasonable to assume that capacities can be estimated with little error. This is the case for airports located in areas where there is little variability in weather conditions or for which changes in weather conditions are predictable and weather patterns remain stable for a long time period, once established. In addition, when the problem is extended to include airborne flights as well as those on the ground, the deterministic model will be seen to be applicable.

2.1 Standard Formulations

2.1.1 Mathematical Model

We consider arrival operations at a given destination airport (airport Z) during a time interval $[0, T]$ for which we expect some congestion. The interval $[0, T]$ is subdivided into P consecutive time periods T_1, T_2, \dots, T_P with corresponding fixed deterministic capacities k_1, k_2, \dots, k_P . There are N flights F_1, F_2, \dots, F_N scheduled to land at airport Z during these time periods. For each flight F_i we know the index of the flight's scheduled landing period, denoted by P_i ; and the cost $C_{g_i}(x)$ of delaying flight F_i for x time periods on the ground before take-off. We assume that all flights that were not

able to land during one of the time periods T_1, T_2, \dots, T_P can do so during a final time period T_{P+1} (i.e. we assume that $k_{P+1} = \infty$).

The objective is to find the ground holding policy X_1, X_2, \dots, X_N , where X_i is the number of time periods flight F_i is delayed on the ground. The ground holding policy must be feasible (i.e. it does not violate the capacities K_i) and minimize the total ground delay cost:

$$TC = \sum_{i=1}^N C_{gi}(X_i)$$

The underlying assumptions of this model are as follows:

- (1) Imposing a ground delay X_i on flight F_i will make that flight arrive at airport Z during time period $T_{P_i+X_i}$. This requires that travel times be deterministic and that airport Z be the only congested element of the air traffic network (i.e. no delays can occur at the origin airports or en route to airport Z).
- (2) We can determine in advance how available capacity at airport Z will be allocated between arrivals and departures.

Both assumptions can in many cases be reasonably realistic given the nature of the model. The assumption that airport Z is the only congested element of the network is the key one in virtually all the problems this report is concerned with. As was stated in the introduction, the approach taken in this report is to decompose the problem by looking at the single airport GHPP as a first approximation.

We define the assignment variables x_{ij} by $x_{ij} = 1$ if flight F_i is assigned to land during period T_j , $x_{ij} = 0$ otherwise. The x_{ij} 's are only defined for $j \geq P_i$.

We denote by C_{ij} the quantity $Cg_i(j-P_i)$, the cost of assigning flight F_i to land during time period T_j . This quantity is also only defined for $j \geq P_i$.

Using this notation the solution of the following integer program (IP), hereafter called **Problem I**, yields the optimal policy:

Problem I

min

$$\sum_{i=1}^N \sum_{j=P_i}^{P+1} C_{ij} x_{ij} \quad (1)$$

subject to

$$\sum_{j=P_i}^{P+1} x_{ij} = 1 \text{ for all } i = 1, \dots, N \quad (2)$$

$$\sum_{i=1}^N x_{ij} \leq k_j \text{ for all } j = 1, \dots, P \quad (3)$$

$$x_{ij} = 0 \text{ or } 1 \text{ for all } (i,j) \quad (4)$$

The constraint matrix in Problem I is totally-unimodular. We could therefore relax the integrality conditions and use the simplex method to solve it. In fact, as we will see in the next section, the solution of Problem I corresponds to a minimum cost flow in a capacitated network and we can therefore use faster specialized algorithms. We will also see that a slight reformulation of the network approach poses the problem as a classical assignment problem for which more specialized algorithms exist. (Ahuja [1988] is an excellent reference for network optimization algorithms.) Finally, we will see that

an even faster algorithm can be developed when the cost functions $Cg_i(X_i)$ satisfy certain conditions that we will identify in section 2.2.

2.1.2 Minimum Cost Flow

Figure 2-1 illustrates a capacitated network formulation of Problem I for the case of general costs functions $Cg_i(t)$.

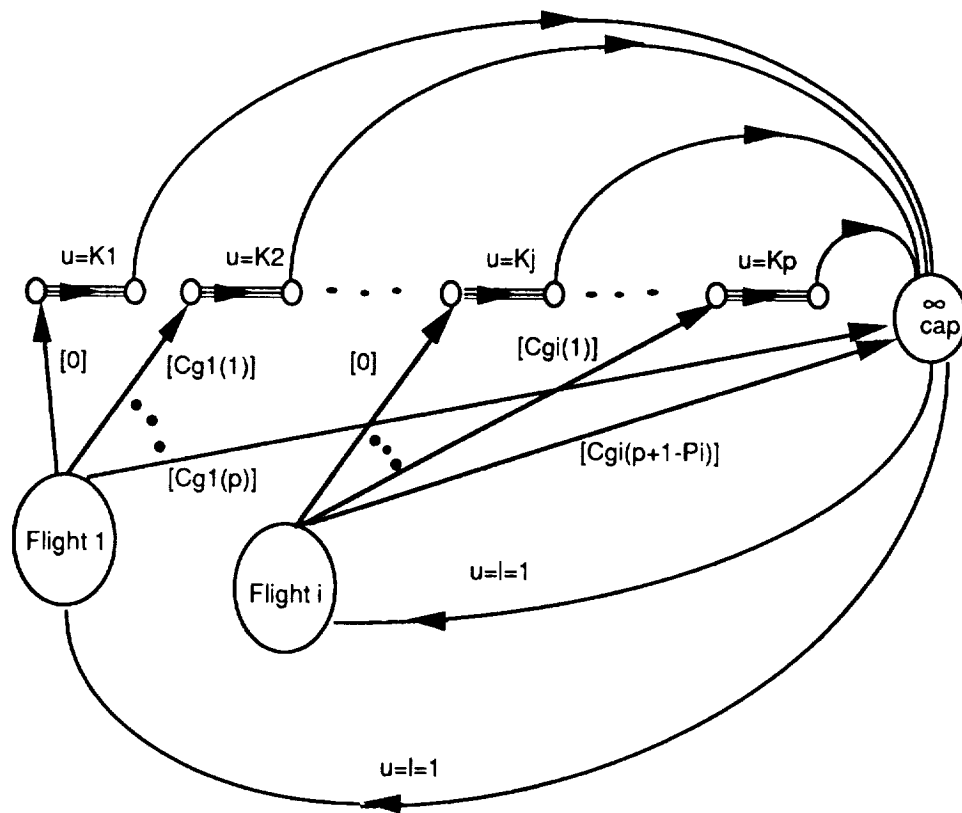


Figure 2-1

The numbers in brackets represent the costs per flow-unit associated with the corresponding arc. When no number is indicated this cost is assumed to be zero.

The letters u and l represent respectively the upper and lower limits for the flow on each arc. The default values are infinity for the upper bound and zero for the lower bound.

Each time period T_i is represented by an arc with the upper bound for the flow on this arc set to the capacity of the time period ($u=k_i$). Each flight F_i generates a node “flight i ” in the above network. Each node “flight i ” is connected to the nodes at the origin of the arcs representing all time periods with index $\geq P_i$.

There are two different ways to put a limit on the length of individual ground holds. The first is to have a cost structure with rapidly increasing marginal costs. (For example: a large α in the cost function of Section 2.2.1) The second is to put a constraint directly onto the ground holds.

If it is desired that ground holds have an upper bound $= d$, then arcs from node “flight i ” would only go to nodes representing time periods with indices between P_i and $(P_i + d)$. However, now there would have to be some provision for checking to see if there is a feasible solution.

It is clear that the cost of any feasible flow through the network is:

$$\sum_{i=1}^N C_{gi}(j - P_i)$$

where j is the index of the time period corresponding to the non-zero flow out of the node "flight i ".

The optimal assignment therefore corresponds to the minimum cost flow through this network.

2.1.3. Assignment Problem

We can also consider minimum separations between landing aircraft explicitly. Instead of considering the period arcs with upper bounds K_j to correspond to time intervals, with each interval having the capacity to serve several flights, we can view them as individual landing slots for single aircraft. The number of landing slots to be generated for each time period is equal to the estimated capacity for that time period. The problem becomes a classical assignment problem, as illustrated by figure 2-2.

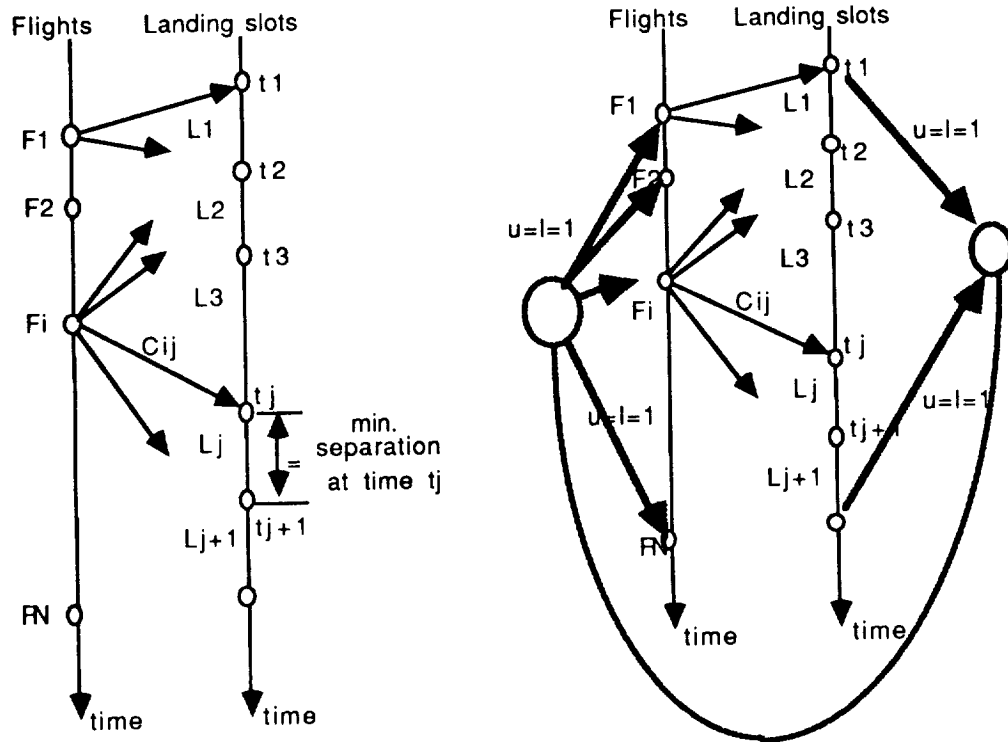


Figure 2-2

If S_i is the scheduled arrival time for flight F_i then the cost C_{ij} of assigning flight F_i to landing slot L_j is given by: $C_{ij} = C_{gi}(t_j - S_i)$ if $t_j \geq S_i$

This can also be solved as a minimum cost flow problem if we introduce some dummy nodes and arcs as shown in figure 2-2. But to be able to use faster algorithms tailored to the assignment problem we have to make sure that the cumulative number of available landing slots does not exceed the number of flights scheduled at any point in time and that the cost matrix is complete. We accomplish this in the following manner:

Never create a landing slot at time t if the cumulative number of scheduled arrivals at that time is lower than the cumulative number of already created landing slots. Doing

this ensures that the number of available assignments equals the number of flights and still produces an optimal solution since it is clear that no assignment will use more landing slots than the flights can fill according to the earliest schedule. We also need to complete the cost matrix; if we index the times for the landing slots as t_1, t_2, \dots, t_N we can do so by setting $C_{ij} = C_{gi}(t_N - S_i) + 1$, for $t_j < S_i$.

If it is desired that ground holds have an upper bound, then slots corresponding to holds which are too long could be assigned arbitrarily large costs. A solution with very large total cost would then tip off the presence of an infeasible solution (infeasible meaning ground holds larger than the upper bound).

The time complexity of the better algorithms to solve the assignment problems is roughly $O(N^{2.5})$ to $O(N^3)$ where N is the number of flights (or equivalently the number of landing slots). Since we are typically dealing with time periods of several hours and the number of (arriving) flights as high as several hundred or even a thousand, it is reasonable to try to develop much faster algorithms specifically tailored to our problem. In the following section we will develop such an algorithm that takes advantage of special forms for the cost functions $C_{gi}(t)$.

2.2 The Fast Algorithm

2.2.1. Regular Costs

We utilize the mathematical formulation developed in section 2.2.1 and denote by $\Delta_j C_{gi}$ the marginal cost of delaying the arrival of flight F_i by one period, from time period T_j to T_{j+1} :

$$\Delta_j C_{gi} = C_{gi}(j+1-P_i) - C_{gi}(j-P_i) \text{ for } j \geq P_i$$

A simple algorithm can be developed when the cost functions $C_{gi}(t)$ satisfy the following two conditions:

For any pair of flights (F_i, F_k) :

(i) if $\Delta_j C_{gi} > \Delta_j C_{gk}$ for some time period T_j , then for *any* $m > j$,

$$\Delta_m C_{gi} > \Delta_m C_{gk}, \text{ and}$$

(ii) if $\Delta_j C_{gi} = \Delta_j C_{gk}$ for some time period T_j , then for *any* $m > j$,

$$\Delta_m C_{gi} = \Delta_m C_{gk}.$$

Intuitively, these conditions say that if it is more costly to delay flight F_i than flight F_k during time period T_j , then it never becomes cheaper to delay flight F_i rather than F_k later on. In following discussions we will refer to conditions (i) & (ii) as *regularity conditions* and costs which satisfy the regularity conditions as “regular costs.”

Let b_i be the cost of delaying flight F_i for one time period on the ground and x be the number of time periods delayed. Then, an interesting class of cost functions is given by the following:

$$C_{gi}(x) = \begin{cases} \frac{b_i(1+\alpha)^x - 1}{\alpha} & \text{if } \alpha \neq 0 \text{ and} \\ b_i x & \text{if } \alpha = 0 \end{cases}$$

This results in marginal cost functions given by: $\Delta_j C_{gi} = b_i(1+\alpha)^{j-P_i}$ for $j \geq P_i$.

These cost functions satisfy the regularity conditions, seen as follows:

Assume that, for two given flights F_i and F_k , we have $\Delta_j C_{gi} > \Delta_j C_{gk}$ for some time period T_j . This translates into $b_i(1+\alpha)^{j-P_i} > b_k(1+\alpha)^{j-P_k}$. Now, for any $m > j$, we can multiply both sides of the inequality by $(1+\alpha)^{m-j}$ yielding:

$$b_i(1+\alpha)^{m-P_i} > b_k(1+\alpha)^{m-P_k} \text{ or } \Delta_m C_{gi} > \Delta_m C_{gk}.$$

The magnitude of α affects the distribution of delays among classes of aircraft. Setting α to 0 corresponds to assuming linear cost functions which tends to assign a disproportionate amount of delays to smaller aircraft; on the other hand, a very high α will result in assigning aircraft to available capacity on a first come first served basis (as is usually done under present practice). Note that

$$\alpha = \frac{\Delta_j C_{gi} - \Delta_{j-1} C_{gi}}{\Delta_{j-1} C_{gi}}, \text{ and so } \alpha \text{ can be interpreted as the relative increase in$$

cost due to holding a flight on the ground for an additional hour.

2.2.2. Notation, Pseudo-Code and a Flow-Chart

The following notation will be used in the algorithm that can be used when the regularity conditions apply:

E_j = set of indices of flights eligible for operation during time period T_j under the optimal policy.

OP_j = set of indices of flights that actually operate during time period T_j under the optimal policy.

The following algorithm produces an optimal solution to Problem I, if the costs satisfy the regularity conditions; see Terrab [1990] for details. This algorithm (called the “Fast Algorithm”) determines the optimal policy X_1, X_2, \dots, X_N and has worst-case

complexity $O(P N \ln N)$ due to the sorting steps. Throughout the algorithm, i is one of the flight indices (between 1 and N inclusive) and j is one of the landing time period indices (between 1 and P inclusive).

procedure Fast Algorithm

begin

$E_0 = \emptyset$

$OP_0 = \emptyset$

for $j = 1$ to P **do**

begin

$E_j = \{ i : P_i = j \} \cup \{ i : i \text{ in } E_{j-1}, i \text{ not in } OP_{j-1} \}$

$OP_j = \{ i : i \text{ in } E_j, i \text{ the index of one of the } k_j \text{ highest } \Delta_j C_{gi} \}$

end do

for $i = 1$ to N **do**

begin

if for some j there is an i in OP_j **then**

$X_i = j - P_i$

else

$X_i = P + 1 - P_i$

end if

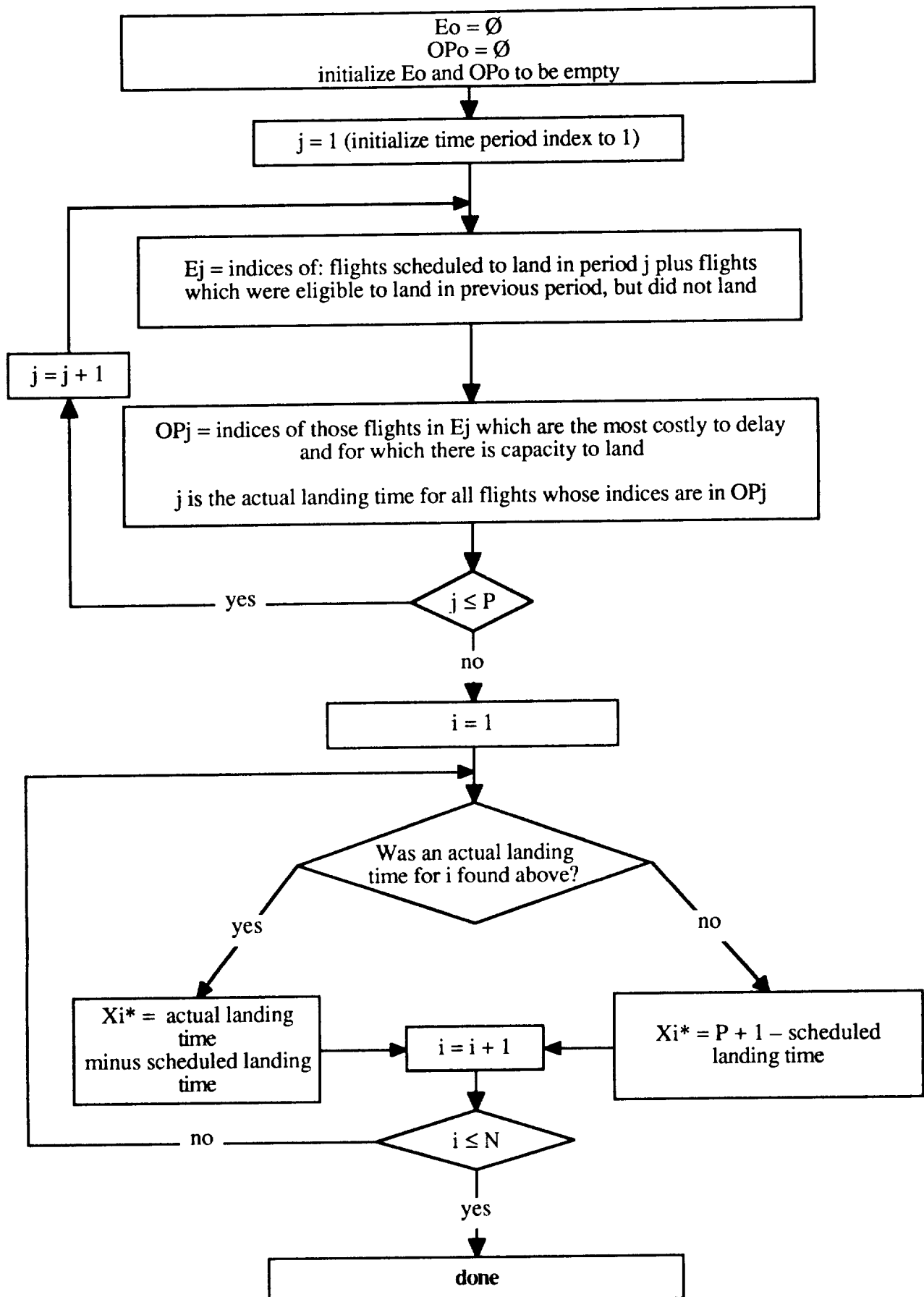
end do

end procedure

The basic idea of the algorithm is to order candidate flights for landing according to their marginal cost of delay for each time period T_j and allow the k_j flights with highest costs to land.

Control over the maximum length of ground holds again could be introduced either by using an appropriate cost structure (e.g. large α) or limiting ground holds directly by including flights in the OP_j sets in such a way that the upper bound on ground holds is not violated. That is, a flight which is at the ground hold limit would be go to the “head of the line.” There is no guarantee of optimality (or feasibility) in this problem, with bounds on ground holds.

The following is a flow chart of the Fast Algorithm.



2.2.3. A Simple Example

An example will help illustrate the Fast Algorithm. Suppose we are given the following deterministic information:

Flight	Marginal cost for holding (per period)	Capacity	Scheduled landing period
1	10	2	3
2	5	2	5
3	10	2	5
4	10	2	4
5	10	2	3
6	5	2	5
7	5	2	3
8	10	2	5
9	10	2	4
10	5	2	4

The Fast Algorithm:

Initialization:

$$E_0 = \emptyset$$

$$OP_0 = \emptyset$$

Since there are no flights scheduled to land before period 3, there are no eligible flights and hence no assigned flights in the first two periods:

$$E_1 = \emptyset$$

$$OP_1 = \emptyset$$

$$E_2 = \emptyset$$

$$OP_2 = \emptyset$$

In period 3, because there were no previous flights unable to land, the eligible flights are those scheduled to land (1, 5 and 7) ; thus

$$E_3 = \{ 1, 5, 7 \}$$

and since the two (capacity) highest marginal costs are associated with flights 1 and 5 we have:

$$OP_3 = \{ 1, 5 \}.$$

In period 4, the eligible flights are those scheduled to land (4, 9 and 10) along with any eligible in the previous period which did not land (flight 7); therefore

$$E_4 = \{ 4, 9, 10 \} \cup \{ 7 \} = \{ 4, 7, 9, 10 \}$$

and since the two (capacity) highest marginal costs are associated with flights 4 and 9 we have:

$$OP_4 = \{ 4, 9 \}.$$

In period 5, the eligible flights are those scheduled to land (2, 3, 6 and 8) as well as those eligible in the previous period which did not land (flights 7 and 10); thus

$$E_5 = \{ 2, 3, 6, 8 \} \cup \{ 7, 10 \} = \{ 2, 3, 6, 7, 8, 10 \}$$

and since the two (capacity) highest marginal costs are associated with flights 3 and 8 we have:

$$OP_5 = \{ 3, 8 \}.$$

In period 6, there are no flights scheduled to land, so the only eligible flights are those eligible in the previous period which did not land (flights 2, 6, 7 and 10); therefore

$$E_6 = \emptyset \cup \{ 2, 6, 7, 10 \} = \{ 2, 6, 7, 10 \}.$$

All the eligible flights have the same marginal cost (\$5), to break the tie we give priority to the flights which have been delayed the longest; since capacity is two, this results in:

$$OP_6 = \{ 7, 10 \}.$$

In period 7, there are no flights scheduled to land, so the only eligible flights are those eligible in the previous period which did not land (flights 2 and 6); thus

$$E_7 = \emptyset \cup \{ 2, 6 \} = \{ 2, 6 \}.$$

There is capacity for both flights; therefore

$$OP_7 = \{ 2, 6 \}.$$

From the information in the OP_i sets, we know the actual landing times. For example, $OP_7 = \{ 2, 6 \}$ tells us that flights 2 and 6 are actually going to land in period 7. The difference between the actual landing times and the scheduled landing times are the values of the ground holds, so we get the following:

Flight	Marginal cost for holding (per period)	Capacity	Scheduled landing period	Actual landing period	Ground hold
1	10	2	3	3	0
2	5	2	5	7	2
3	10	2	5	5	0
4	10	2	4	4	0
5	10	2	3	3	0
6	5	2	5	7	2
7	5	2	3	6	3
8	10	2	5	5	0
9	10	2	4	4	0
10	5	2	4	6	2

There are a total of 9 flight-periods of ground holds, all costing \$5 per period; the total cost is therefore \$45.

It is possible to modify the example by introducing a maximum time for ground holds. Revising the algorithm is easy; but there is no guarantee now that a solution produced by the algorithm is optimal for the modified problem.

For example, suppose ground holds must be no more than two periods; then flight 7 violates this constraint in the solution above. To meet this constraint we could do the following: when flight 7 (already holding for 2 periods) is eligible in period 5, assign it precedence over the other flights. Then the resulting solution is different from the previous one; here $OP_5 = \{ 3, 7 \}$ and $OP_6 = \{ 8, 10 \}$ and the total cost of the solution is \$50. This is most likely the minimum cost for this example; but in general, an optimal policy and associated minimum cost for the modified problem are not known.

2.3 Adaptive Considerations

Since we have assumed that everything is known and deterministic, there are no airborne delays in the model; all delays are taken on the ground. However, since the Fast Algorithm can be used in real-time ($O(P N \ln N)$) we can run it at appropriate intervals during the day, initialized with the current state of the system. This could include flights which are airborne as well as those on the ground; the costs would have to reflect the various states of the flights. Thus the “ground hold” (in the model) costs of the airborne flights would be actually be airborne hold costs. As long as the regularity conditions were satisfied, the Fast Algorithm provides optimal solutions. For general cost function, the assignment problem (with algorithms $O(N^{2.5})$) could be used to find optimal solutions.

3. PROBABILISTIC CASE

3.1 Problem Description

We consider landing operations at an arrival airport Z during consecutive time periods T_1, T_2, \dots, T_P . The capacity of the airport during time period T_j is given by K_j , now modeled as a random variable rather than a fixed known quantity. This is the only place where stochastic behavior is modeled; all other assumptions are the same as in the deterministic case. There are N flights, F_1, F_2, \dots, F_N , scheduled to land during these time periods. For each flight we define:

- X_i = the ground hold time for flight F_i .
- $C_{gi}(x)$ = the cost of delaying flight F_i if $X_i = x$
- $C_{aj}(x, y)$ = the cost of delaying flight F_i for y time periods in the air when it has already been delayed x time periods on the ground (for a total delay of $x+y$ time periods).
- P_i = the index of the earliest possible landing time period (scheduled arrival) for flight F_i .
- Π_i = the priority of flight F_i . A “fixed priority rule” is defined as follows: if two flights F_i and F_j are candidates for landing during the same time period, F_i will not be cleared for landing before F_j if $\Pi_i < \Pi_j$. Note that the assumption of a fixed priority rule is not particularly restrictive in practice; air traffic controllers tend to use such rules in practice when sequencing flights for landing. Priority by scheduled time of arrival is an example of a landing priority rule used in practice. In addition, if we

use (ground and airborne) delay cost functions that satisfy the regularity conditions then we can impose such a fixed priority rule with no loss of optimality. In the rest of the report, we assume that the flights have been reordered so that $\Pi_{i+1} > \Pi_i$ for all i .

- \mathbf{H}_i = $(H_i^1, H_i^2, \dots, H_i^P)$ where H_i^j represents the number of flights with priority greater than Π_i that are assigned to arrive during time-period T_j .

- C_i = the (now ground plus air) delay cost; this is a random variable because of the random capacities K_1, K_2, \dots, K_P

An important result is that $E[C_i]$ can be written as a function of two quantities, X_i and \mathbf{H}_i ; given $C_{gi}(x)$, $C_{ai}(x, y)$, P_i , and Π_i , and the probability distribution of the random capacities K_1, K_2, \dots, K_P . In order to show this, some additional notation is needed.

Suppose that there are only C different capacity “scenarios”, where an individual scenario is defined as a sequence of (observed) values of the random capacities, K_1, K_2, \dots, K_P .

(Observed values are the values random variables can take on for a given realization; they are simply numbers.) Define the following for scenario j (out of the C scenarios):

- k_m^j = the (observed) value of the random capacity K_m in scenario j

- p_j = the probability of scenario j occurring

- w_i^j = index of the actual landing time period for F_i

- L_h^j = the number of flights with higher priority than F_i that are still waiting for landing at the beginning of time period T_h . L_h^j can be calculated using the following recursion:

- 1) $L_0^j = 0$

- 2) **for** $h = 1$ to P **do** $L_h^j = \max\{H_i^h + L_{h-1}^j, 0\}$

Then $w_i^j = \min\{ m : m \geq P_i + X_i \text{ and } H_i^m + L_{m-1}^j < k_m^j \}$. These relationships express the fact that flight F_i will only be able to land during the first time period after $T_{P_i+X_i}$, $T_{w_i^j}$, for which the cumulative capacity is such that all flights with priority higher than Π_i that are scheduled before $T_{w_i^j}$ have been able to land.

- Y_i^j = airborne delay for flight F_i . Y_i^j is given by the formula:

$$Y_i^j = w_i^j - P_i - X_i.$$

Note that L_h^j , w_i^j , and hence Y_i^j are calculated using only H_i , for given P_i and X_i . Since we can write

$$E[C_i] = C_{g_i}(X_i) + \sum_{j=1}^C C_{a_i}(X_i, Y_i^j) p_j \equiv c_i(X_i, H_i)$$

we have shown that $E[C_i]$ can be written as a function of X_i and H_i ; we will use the notation $E[C_i] = c_i(X_i, H_i)$. The probabilistic version of Problem I is to minimize $\sum_{i=1}^N E[C_i]$.

An exact optimal solution to the probabilistic GHPP can be found through a dynamic programming approach, if it is assumed that a fixed landing priority rule $\Pi_1, \Pi_2, \dots, \Pi_N$ is imposed on the flights F_1, F_2, \dots, F_N . The time complexity for running this dynamic program (DP) is $O[C(P+1)^2((N/P)+1)^P]$, thus the approach is not practical in a realistic scenario. See Terrab[1990] for details.

3.2 The MMR Algorithm

3.2.1. Cost Functions, Pseudo-Code and a Flow-Chart

A fast algorithm was developed which can be interpreted as a maximum marginal return (MMR) algorithm under suitable conditions; hence, it will be called “the MMR Algorithm.”

If we use (ground and air-delay) cost functions which satisfy the regularity conditions and use the resulting landing priority rule, the MMR Algorithm translates into a strategy consisting of going after the highest possible marginal cost reduction at each step of the algorithm. This is because the landing priority rule associated with a set of regular cost functions is such that the flights with highest delay costs are given priority; therefore the MMR Algorithm goes after the highest possible gains from the outset.

The MMR Algorithm works with general cost functions; however there are some interesting cost functions which satisfy the regularity conditions. We will assume that $Ca_i(x,y)$, the cost of holding flight F_i for y time periods in the air, if it has been held x time periods on the ground before take-off, is similar to the cost function in the deterministic case so that the marginal cost of delaying flight F_i during time period T_j (where $j \geq P_i + x$) in the air, denoted $\Delta_j Ca_i$, is given by:

$$\Delta_j Ca_i \equiv Ca_i(x, j - P_i - x + 1) - Ca_i(x, j - P_i - x) = \mathcal{K} b_i (1 + \alpha)^{x-1} (1 + \beta)^{j - P_i - x + 1}$$

When we assume that $\alpha = \beta$ (as is done in the next section) we get $\Delta_j Ca_i = \mathcal{K} b_i (1 + \beta)^{j - P_i}$; in which case \mathcal{K} can be interpreted as a multiplicative coefficient intended to reflect the ratio between the direct airborne operating costs of aircraft and the direct costs of keeping

them on the ground. The coefficient β plays, for airborne delay costs, the same role α plays for ground delay costs. We note that

$$\beta = \frac{\Delta_j C_{ai} - \Delta_{j-1} C_{ai}}{\Delta_{j-1} C_{ai}}, \text{ which is the } \underline{\text{relative increase in cost}} \text{ due to holding a}$$

flight in the air for an additional time period.

The algorithm assumes that a fixed landing priority rule is in effect. The basic idea for it is motivated by the result found in the previous section: if we have a fixed landing priority rule, $E[C_i]$ (the expected cost for flight F_i) depends only on X_i and the status of flights of higher priority represented by the vector H_i , and not on flights with lower priority. (Again: H_i^j , the j^{th} component of H_i , represents the number of flights with priority greater than Π_i that are assigned to arrive during time-period T_j .) In the algorithm, H_{i-1} is calculated as a function of H_i and X_i as follows:

- For time period $T_{P_i+X_i}$ to which flight F_i is reassigned set

$$H_{i-1}^{P_i+X_i} = H_i^{P_i+X_i} + 1.$$

This expresses the fact that the number of flights with higher priority than flight F_{i-1} scheduled during time period $T_{P_i+X_i}$ has to be increased by one unit if flight F_i is reassigned to that time period.

- For all other time periods T_r with $r \neq P_i + X_i$ set

$$H_{i-1}^r = H_i^r,$$

since the number of flights with higher priority than flight F_{i-1} scheduled during time period T_r does not change if F_i is reassigned to another time period.

Remember that flights have been indexed so that $\Pi_{i+1} > \Pi_i$ for all i . Thus, the flight with highest priority, F_N , is such that $\mathbf{H}_N = \mathbf{0}$. Therefore, the expected (ground+air delay) cost for flight F_N , $E[C_N] = c_N(X_N, \mathbf{0})$, does not depend on the status of any other flight; it depends only on the ground hold X_N that we impose on it. Thus, we can find the ground hold X_N^* that leads to the lowest expected delay cost for F_N . Once we have X_N^* , we can compute $\mathbf{H}_{N-1}^*[0, X_N^*]$ as described above and we can therefore also compute, for each possible ground hold X_{N-1} ,

$$E[C_{N-1}] = c_{N-1}(X_{N-1}, \mathbf{H}_{N-1}^*[0, X_N^*])$$

the expected cost for flight F_{N-1} . Again this allows us to find the ground hold X_{N-1}^* for flight F_{N-1} as the one that yields the lowest expected cost. This procedure is repeated until we have computed the ground hold for the flight with lowest landing priority, F_1 . At the end, the total expected cost can be calculated by summing the individual $E[C_i]$'s.

The algorithm therefore computes ground holds by minimizing the expected cost for each flight individually, starting with the flight with highest priority. We note that a major component of this algorithm is the computation, for a given flight F_i , of the expected cost associated with a given ground hold, X_i , and a given vector \mathbf{H}_i . We saw previously that this computation consists of looking at each of the C capacity scenarios to find the actual landing time (and therefore the airborne delay) for flight F_i corresponding to that capacity scenario. This involves a search through, at most, $P+1$ time periods. The time complexity of this procedure is therefore $O[NC(P+1)^2]$.

Pseudo-code follows:

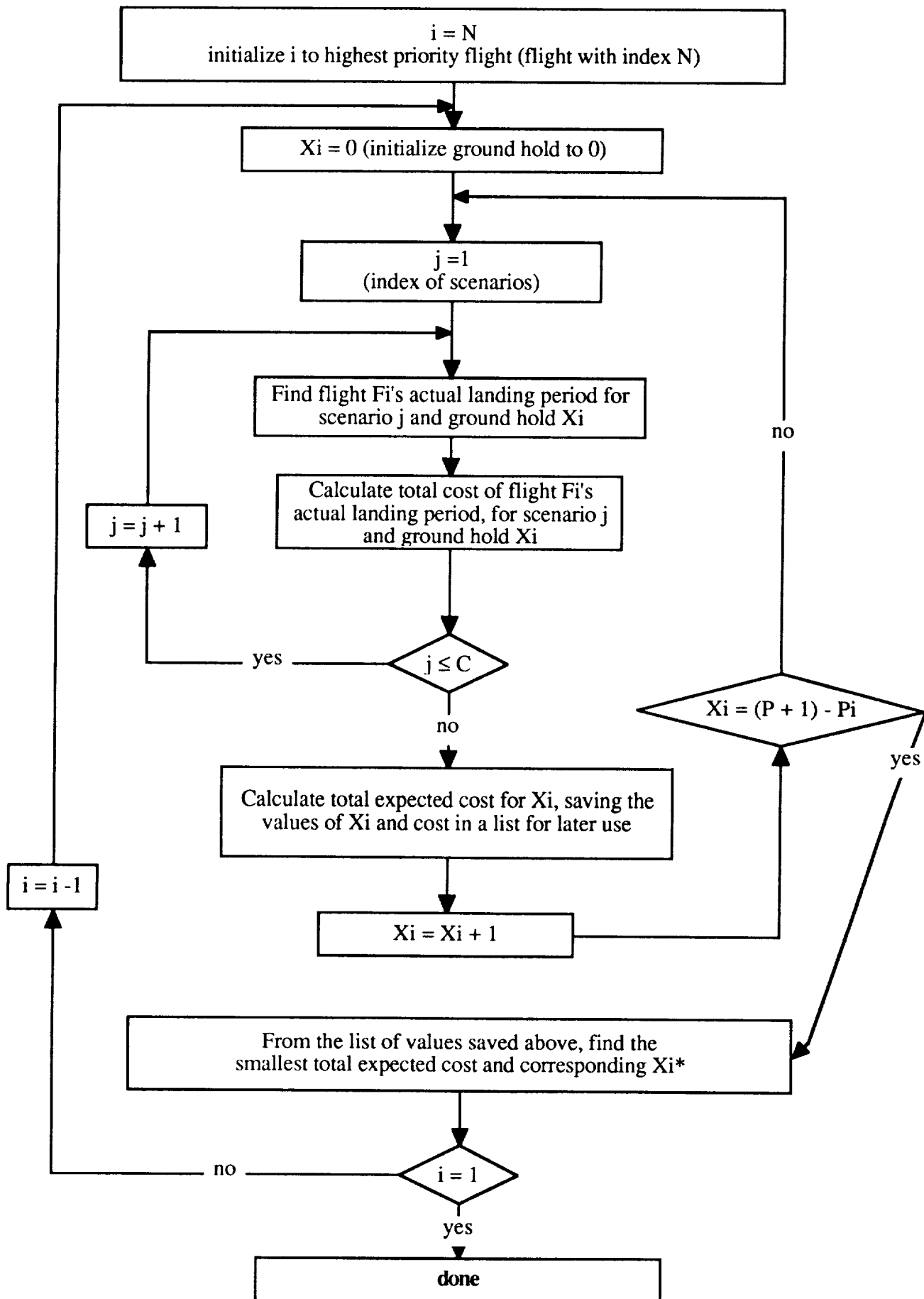

```

procedure MMR Algorithm
begin
    let  $H_N = 0$ 
    for  $i = N$  downto 1 do
        begin
            find  $X_i^*$  such that  $c_i(X_i^*, H_i) = \min \{ c_i(X_i, H_i) \}$ 
            for  $r = 1$  to  $P$  do
                begin
                    if  $r = P_i + X_i^*$  then
                         $H_{i-1}^r = H_i^r + 1$ 
                    else
                         $H_{i-1}^r = H_i^r$ 
                    end if
                end do  $r$ 
            end do  $i$ 
        end procedure MMR Algorithm

```

Just as in the deterministic case, a limit on the ground holds could be implemented through the cost functions (e.g. large α and β in the cost functions of Section 3.2.1) or directly. To do it directly is very straightforward; the selected ground hold X_i^* would be the one which minimizes the expected cost c_i among those X_i less than or equal to the upper bound on the ground holds.

The following is a flow chart of the MMR Algorithm.



3.2.2. A Simple Example

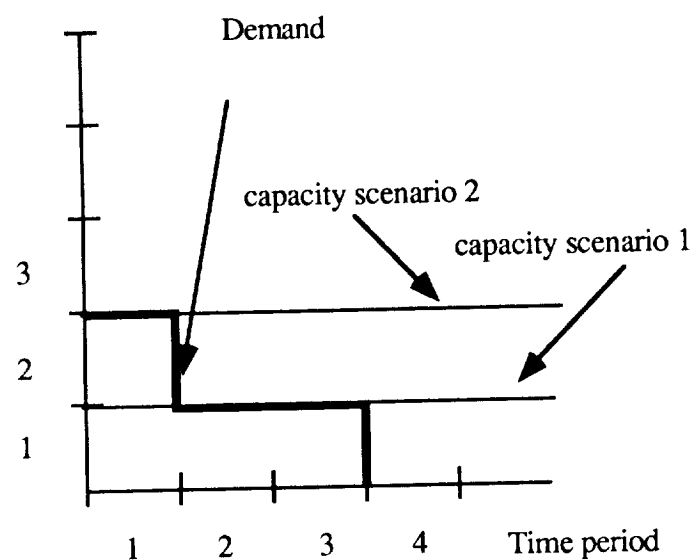
An example will help illustrate the MMR Algorithm; remember that flights are numbered in increasing priority. Suppose we are given the following probabilistic information:

Flight	Marginal cost per period		Scheduled landing period
	ground holds	airborne holds	
1	10	20	3
2	10	20	1
3	5	10	1
4	10	20	2

Assume that there are two capacity scenarios;

scenario 1: capacity is 1 for all periods, with probability 0.4

scenario 2: capacity is 2 for all periods, with probability 0.6.



The MMR Algorithm:

Step i = 4:

Since there are no flights with higher priority than flight 4, we have:

$$H_4 = (0, 0, 0, 0)$$

$$X_4^* = 0 \quad (\text{no ground hold needs to be considered for flight 4; no capacity problem yet})$$

$$H_3 = (0, 1, 0, 0) \quad (\text{since flight 4 assigned to land in period 2}).$$

Flight 4 can land when scheduled with no holds in either scenario; total cost is 0.

Step i = 3:

We find:

$$X_3^* = 0 \quad (\text{no ground hold needs to be considered for flight 3; no capacity problem yet})$$

$$H_2 = (1, 1, 0, 0) \quad (\text{since flight 3 assigned to land in period 1}).$$

Flight 3 can land when scheduled with no holds in either scenario; total cost is 0.

Step i = 2:

There are three ground hold values which need to be examined: 0, 1 or 2. ($X_2 > 2$ has higher cost than $X_2 = 2$ with probability one.) The following two tables contain the costs incurred from each of these values, for the given scenario.

		Scenario 1 (with probability 0.4)			
		$X_2 \rightarrow$	0	1	2
costs	ground		0	10	20
	airborne		40	20	0
	total		40	30	20

		Scenario 2 (with probability 0.6)			
		$X_2 \rightarrow$	0	1	2
costs	ground		0	10	20
	airborne		0	0	0
	total		0	10	20

The expected total cost, as a function of the ground holds is:

Ground hold	Expected total cost
0	$0.4 (40) + 0.6 (0) = 16$ * minimum
1	$0.4 (30) + 0.6 (10) = 18$
2	$0.4 (20) + 0.6 (20) = 20$

Thus $X_2^* = 0$ and $H_1 = (2, 1, 0, 0)$ since flight 2 is assigned to land in period 1.

Flight 2 will either be able to land when scheduled (in period 1) with probability 0.6 (under scenario 2) or will be delayed in the air two periods with probability 0.4 (under scenario 1); flight 2 contributes \$16 to total expected costs.

Step i = 1

There are two ground hold values which need to be examined: 0 or 1. ($X_2 > 1$ has higher cost than $X_2 = 1$ with probability one.) The following two tables contain the costs incurred from each of these values, for the given scenario.

		Scenario 1 (with probability 0.4)	
		$X_1 \rightarrow$	
		0	1
costs	ground	0	10
	airborne	20	0
	total	20	10

		Scenario 2 (with probability 0.6)	
		$X_1 \rightarrow$	
		0	1
costs	ground	0	10
	airborne	0	0
	total	0	10

Thus the expected total cost, as a function of the ground holds is:

Ground hold	Expected total cost	
0	$0.4 (20) + 0.6 (0) = 5$	* minimum
1	$0.4 (10) + 0.6 (10) = 10$	

Thus $X_1^* = 0$ since flight 1 is assigned to land in period 3.

Flight 1 will either be able to land when scheduled (in period 3) with probability 0.6 (under scenario 2) or will be delayed in the air one period with probability 0.4 (under scenario 1); flight 1 contributes \$ 5 to total expected costs.

Assignments are as follows:

Flight	Ground hold	Scheduled to land in period	Assigned to land in period
1	0	3	3
2	0	1	1
3	0	1	1
4	0	2	2

and the total expected cost is

$$\$0 + \$0 + \$16 + \$5 = \$21.$$

Suppose the probabilities had been reversed, so that scenario 1 has probability = 0.6 and scenario 2 has probability = 0.4. The ground holds for flights 3 and 4 would be the same, but for 1 and 2 we have a different result.

Step i = 2

		Scenario 1 (with probability 0.6)			
		X ₂ →	0	1	2
costs	ground		0	10	20
	airborne		40	20	0
	total		40	30	20

		Scenario 2 (with probability 0.4)			
		X ₂ →	0	1	2
costs	ground		0	10	20
	airborne		0	0	0
	total		0	10	20

Thus the expected total cost, as a function of the ground holds is:

Ground hold	Expected total cost
0	$0.6 (40) + 0.4 (0) = 24$
1	$0.6 (30) + 0.4 (10) = 22$
2	$0.6 (20) + 0.4 (20) = 20$ * minimum

Thus $X_2^* = 2$ and $H_2 = (1, 1, 1, 0)$ since flight 2 is now assigned to land in period 3.

Flight 2 will be able to land in period 3 under either scenario; flight 2 is expected to (and in fact will) incur \$20 in costs.

Step i = 1

There are two ground hold values which are possible: 0 or 1. The following two tables contain the costs incurred from each of these values, for the given scenario.

		Scenario 1 (with probability 0.6)	
		$X_1 \rightarrow$	
		0	1
costs	ground	0	10
	airborne	20	0
	total	20	10
		Scenario 2 (with probability 0.4)	
		$X_1 \rightarrow$	
		0	1
costs	ground	0	10
	airborne	0	0
	total	0	10

Thus the expected total cost, as a function of the ground holds is:

Ground hold	Expected total cost
0	$0.6 (20) + 0.4 (0) = 12$
1	$0.6 (10) + 0.4 (10) = 10$ * minimum

Thus $X_1^* = 1$ and flight 1 will be able to land in period 4 under either scenario; flight 1 is expected to (and in fact will) incur \$10 in costs.

Assignments are as follows:

Flight	Ground hold	Scheduled to land in period	Assigned to land in period
1	1	3	4
2	2	1	3
3	0	1	1
4	0	2	2

and the total expected cost is

$$\$ 0 + \$ 0 + \$ 20 + \$ 10 = \$30.$$

3.3 Adaptive Considerations and Future Directions

The MMR Algorithm ($O[NC(P+1)^2]$) can be used in real time to take advantage of changes in the state of the system. Updated information on the capacities as well as the current state of all flights (on ground, in air, any delays so far, etc.) could be used as the initial values for running the algorithm at appropriate times during the day, perhaps periodically.

Define these two events of interest to airlines:

A = an aircraft is given little or no ground delay and a large airborne delay becomes necessary

B = an aircraft is given a large ground delay when little or none was required

Estimation of the probabilities of these events analytically is non-trivial. However, it would be fairly straightforward to simulate different policies on limiting ground holds (either by increasing the curvature of the cost functions (large α and β values) or directly as constraints) and use the results to make statistical estimates of the probabilities.

The algorithms in this report are a first step to make the tradeoff between ground and airborne delays; further work is needed on the dynamic aspects of the problem. This work would address the estimation issue as well as the following:

- look at the capacity estimation problem
- study general cost functions
- investigate the multi-airport problem
- implementation issues
- study the interaction with tactical flow management actions

4. NUMERICAL EXAMPLES

4.1 Introduction

The computational examples presented below are based on a hypothetical situation constructed as follows:

We consider operations at a given airport Z during a given day according to a schedule that resembles the operation profile at Boston Logan airport during a typical day, using 1987 data. Since 95% of the total daily operations occur between 7-am and 11-pm we restrict our analysis to this time span. The inputs to the problem are:

- A total number of landings for each hour of the day. A random number generator determines the scheduled landing time during each hour for each individual flights using a uniform distribution. (This would be equivalent to simulating the instants of Poisson arrivals, given the number of arrivals per hour.)

- We have 3 types of aircraft distinguished on the basis of their ground delay costs: general aviation, small jets and large jets. The random number generator assigns to each flight one of the three categories according to a prespecified flight mix. For each aircraft type we are given the cost of holding an aircraft of this type one hour on the ground.

4.2 Deterministic Case

The cost function used for the deterministic numerical example was presented in Section 2.2; it was

$$C_{gi}(x) = \begin{cases} \frac{b_i(1+\alpha)^x - 1}{\alpha} & \text{if } \alpha \neq 0 \text{ and} \\ b_i x & \text{if } \alpha = 0 \end{cases}$$

where b_i is the cost of delaying flight F_i for one time period on the ground and α is the relative increase in cost due to holding a flight on the ground for an additional hour.

It is possible to discuss the meaning of the coefficient α at this stage even before looking at the numerical examples. The case $\alpha = 0$ corresponds to linear costs, implying that the cost of holding any flight for a single time period further on the ground does not depend on how long it has been held previously. In this case we expect that an efficient utilization of available capacity will always favor the same type of aircraft, namely the higher cost type to the detriment of less costly types. At the other extreme, the case where α is very large will tend to favor flights on the basis of how long they have already been held on the ground. In fact it is easy to see that for α above a certain threshold value it becomes optimal to assign available capacity to flights on a first-come first-served basis. This happens when α is high enough so that for any pair of aircraft types i, j we find $C_{gi}(2) > C_{gj}(1)$.

Higher values of α correspond to a widespread distribution of ground holds among aircraft types. We can therefore also interpret α as a measure of the distribution of ground holds among aircraft types and indeed use α explicitly as a parameter of the

optimization to distribute these ground holds according to some prespecified criterion. Also, we note again that these cost functions satisfy the regularity conditions and therefore that the Fast Algorithm does yield the optimal solution relative to a particular α .

We assume that a fixed deterministic hourly estimate for the landing capacity at airport Z is available. The following numerical example is intended to illustrate the benefit derived from using the Fast Algorithm. For cost comparison, these results are contrasted with those obtained on the same sample problem from first come, first served (FCFS) policies.

Figure 4-1 shows the relationship of demand to capacity assumed in this numerical example. We assume the (direct operating) costs of holding a flight for one hour on the ground to be \$400 for general aviation, \$1,200 for regular jets and \$2,000 for wide-body jets. The assumed flight mix is 40% general aviation, 40% regular jets, and 20% wide-body jets, similar to that of Logan Airport.

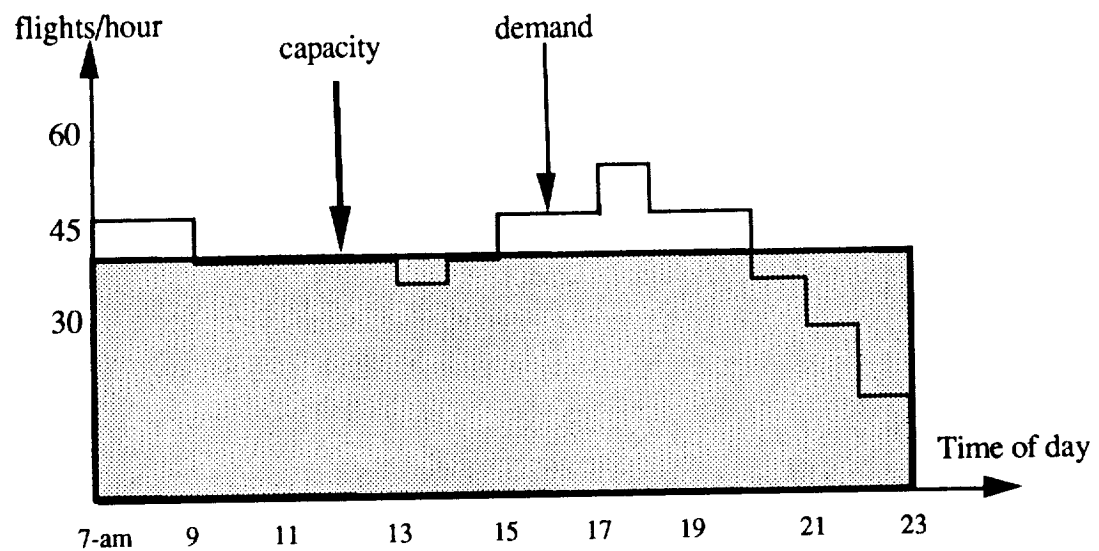


Figure 4-1

Table 4.1 clearly shows the potential for large cost savings from the use of the Fast Algorithm. The trade-off is in average delay per flight. With a first come, first served policy, all flights are treated equally and the average delay times of the three classes of planes are equal (except for chance variation, from the simulation). The Fast Algorithm solution gives average delay times which vary inversely with cost. The most expensive flights are held the least; the least expensive flights are held the most.

Average delay per flight (minutes)		Plane type	Hourly Ground-hold Cost
First come First served	Fast Algorithm Solution		
17.5	28.9	General aviation	\$ 400
16.4	9.7	Standard jets	\$ 1200
17.6	5.9	Wide-body jets	\$ 2000
Total Cost	\$ 191,500		\$ 127,100

•Data from simulated typical day at Logan Airport

Table 4-1

**Large Cost Savings Result from Using the Fast Algorithm
in the Deterministic Case**

4.3 Probabilistic Case

We consider operations at airport Z between 3-pm and midnight with the same cost structure and flight mix (similar to Logan) as in the previous section: 40% general aviation (\$ 400/hour on ground), 40% regular jets (\$ 1200/hour on ground) and 20% wide-body jets (\$ 2000/hour on ground). We will model probabilistic behavior of capacity by assuming three capacity scenarios (for arrivals), given by KAP1 (constant at 30), KAP2 (starting at 40 and dropping 5 at 6PM and again at 9PM) and KAP3 (starting at 40 and increasing to 45 at 7 PM) as pictured in Figure 4-2. The scenarios could be interpreted as: KAP1 representing constant bad weather; KAP2 representing weather getting increasingly worse; and KAP3 representing clearing. The assumed demand profile, based on 1987 Logan Airport data, is shown in dashed lines in the same figure.

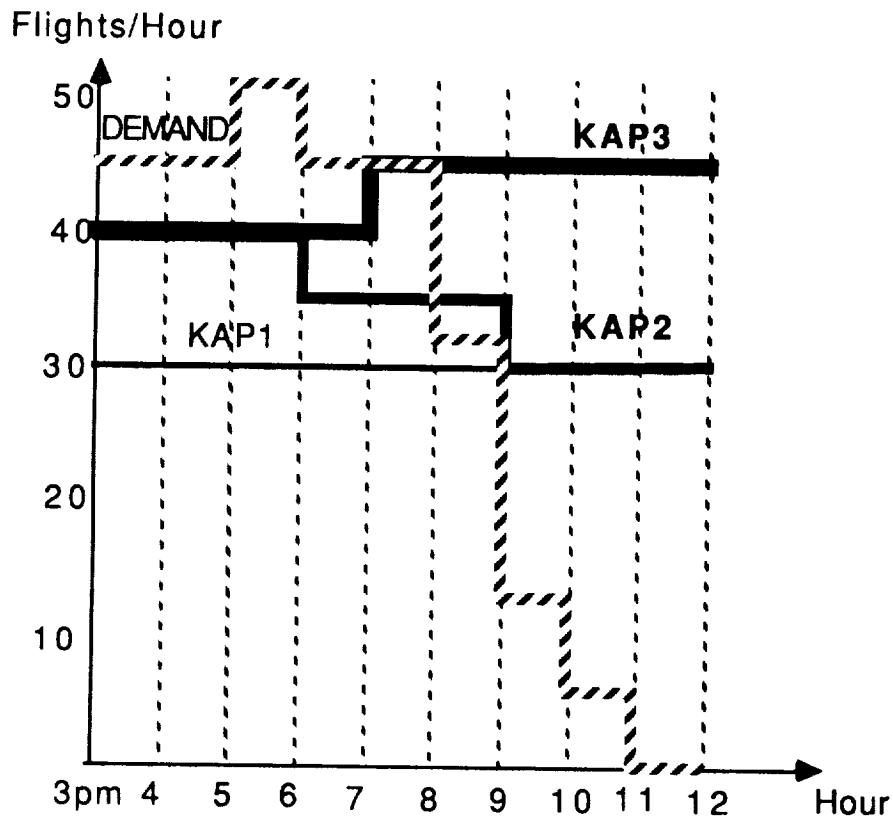


Figure 4-2

We use the expected capacity as the deterministic capacity in a simulation of a first-come, first-served policy (denoted FCFS-T; T for takeoffs). Flights are assigned ground-holds equal to observed airborne delays.

In practice, flights are generally sequenced for landing on a first-come first-served basis (denoted FCFS-L; L for landings). When flights are within 150-200 miles of the destination airport, capacities could be modeled deterministically. Thus the Fast Algorithm could be used on the **airborne** costs to find the optimal landing order. This procedure is denoted “optimal tactics” in Figure 4.3. The cost differentials seen in the table show that studying the tactical problem with a higher fidelity model (one which could be implemented) is clearly warranted.

Data is simulated from a typical day at Logan as described before; airborne costs assumed to be twice ground costs (remember: hourly ground costs are \$400,1200 and 2000 for general aviation, regular jets and wide-body jets respectively).

	Ground costs	Total costs (ground + airborne)	
		with optimal tactics	FCFS-L
MMR	\$ 61,900	\$ 152,300	\$ 326,220
FCFS -T	\$243,300	\$309,220	\$ 415,940
No Holds	\$ 0	\$ 203,520	\$ 529,520

Table 4-2

**Large Cost Savings Result from Using the MMR Algorithm
in the Probabilistic Case**

Many more examples were studied; see Terrab[1990]. Some numerical experiments were designed to show the use of the probabilistic formulations for planning purposes. One of these experiments showed how one can go about evaluating the potential benefit of increasing the reliability of the capacity estimation mechanism in use. The other showed how to compute the marginal benefit of an additional landing slot. Such results can be used to compare the benefits of investments in estimation mechanisms and ATC technologies with their costs.

Also, experimental results showed that the quality of the GHP obtained from optimization methods does not depend on the size of the period length used in the discretization of the time axis (for period length of up to one hour) when we take the uncertainty in travel times into account. This has desirable implications in terms of the complexity of all the algorithms considered in this report .

Another issue likely to determine the practical success of algorithmic approaches is their ability to consider downstream effects. By "downstream" effects we mean situations for which delaying the arrival time of a given aircraft at airport Z has implications that go beyond the direct cost of this delay because, for example, the same aircraft is used for another flight originating at airport Z with a short turnaround time. Situations like these can be modeled through a sharp increase in delay cost beyond a certain level of total (ground+air) delay. The resulting cost functions are likely, however, to violate the regularity conditions and algorithms based on the Fast Algorithm can no longer be used. On the other hand, an min-cost flow or assignment algorithm could be used in the deterministic case and the MMR Algorithm could be used in the probabilistic case in situations without regular cost functions.

The results of this report allow us to feel optimistic about the usefulness of algorithmic approaches to the GHPP. The success of such approaches will ultimately depend on how well they can be integrated into the existing ATC system. It will also depend on their ability to treat users relatively equitably while achieving aggregate cost reductions. Some numerical examples show that significant cost savings occur even when ground and airborne delays are redistributed fairly equitably among classes of aircraft.

5. REFERENCES

Alcabin, M., H. Erzberger, L. Tobias and P.J. O'Brien, *Simulation of Time-Control Procedures for Terminal Area Flow Managment*, **Proceedings of the 1985 American Control Conference**, Boston, MA, June 1985.

Ahuja, Ravindra K., Thomas L. Magnanti and James B. Orlin, *Network Flows*, **Sloan School of Management Working Paper No. 2059-88**, August 1988

Andreatta, G. and G. Romanin-Jacur (1987), *Aircraft Flow Management under Congestion*, **Transportation Science**, 21, 249-253.

Odoni, A.R. (1987), *The Flow Management Problem in Air Traffic Control*, in **Flow Control of Congested Networks** (A.R. Odoni, L. Bianco, G. Szego, editors) Springer-Verlag, Berlin (pp. 269-288).

Terrab, M. (1989), *ATC Flow Control through Ground-Holding*, **Ph.D. Thesis**, Department of Civil Engineering, Mass. Inst. of Technology, Cambridge, MA, February 1990.

Tobias, L., H. Erzberger, H.Q. Lee and P.J. O'Brien, *Mising Four-Dimensional Equipped and Unequipped Aircraft in the Terminal Area*, **Journal of Guidance, Control, and Dynamics**, Vol. 8, 1985, pp. 296-303.

Tobias, L. and J.L. Scoggins, *Time-Based Air traffic Management Using Expert Systems*, **NASA Technical Memorandum 88234**, April 1986, Ames Research Center, Moffett Field, CA.

6. ACKNOWLEDGEMENTS

The authors are grateful to Professor Amedeo Odoni, Co-Director of the OR Center at MIT for his invaluable advice. We also thank Dayl Cohen, Larry Coleman and especially Pike Cook of Massport for assistance in gathering Logan Airport aircraft activity data.

Report Documentation Page

1. Report No. NASA CR-177558		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle Real-Time Adaptive Aircraft Scheduling				5. Report Date June 1990	
				6. Performing Organization Code	
7. Author(s) Stephan E. Kolitz and Mostafa Terrab				8. Performing Organization Report No. A-90198	
				10. Work Unit No. 505-66-11	
9. Performing Organization Name and Address Charles Stark Draper Laboratory, Inc. 555 Technology Square Cambridge, MA 02139				11. Contract or Grant No. NAS2-12419	
				13. Type of Report and Period Covered Contractor Report	
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, DC 20546-0001				14. Sponsoring Agency Code	
15. Supplementary Notes Point of Contact: Leonard Tobias, Ames Research Center, MS 210-9, Moffett Field, CA 94035-1000 (415) 604-5430 or FTS 464-5430					
16. Abstract <p>One of the most important functions of any air traffic management system is the assignment of "ground-holding" times to flights, i.e., the determination of whether and by how much the take-off of a particular aircraft headed for a congested part of the ATC system should be postponed in order to reduce the likelihood and extent of airborne delays. We present an analysis of the fundamental case in which flights from many destinations must be scheduled for arrival at a single congested airport; the formulation is also useful in scheduling the landing of airborne flights within the extended terminal area. We describe a set of approaches for addressing a deterministic and a probabilistic version of this problem. For the deterministic case, where airport capacities are known and fixed, we develop several models with associated low-order polynomial-time algorithms. For general delay cost functions, these algorithms find an optimal solution. Under a particular natural assumption regarding the delay cost function, we develop an extremely fast ($O(n \ln n)$) algorithm. For the probabilistic case, using an estimated probability distribution of airport capacities, we develop a model with associated low-order polynomial-time heuristic algorithm with useful properties.</p>					
17. Key Words (Suggested by Author(s)) Scheduling Air traffic control Assignment problem				18. Distribution Statement Unclassified-Unlimited Subject Category - 03	
19. Security Classif. (of this report) Unclassified		20. Security Classif. (of this page) Unclassified		21. No. of Pages 72	
				22. Price A04	